

Analisi del trasporto ferroviario italiano

Spolaor Andrea – Luglio 2025

Domande

1

Come sono distribuite le stazioni ferroviarie in Italia?

2

Quali sono le regioni che presentano ritardi medi più alti? E quali ritardi minori?

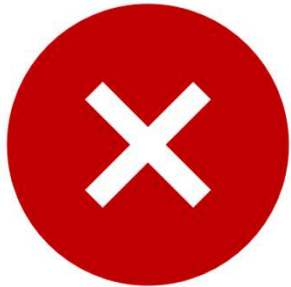
3

Quali sono state le giornate con più cancellazioni? E quelle con più ritardi?

4

Quali sono le tratte con maggior ritardo medio?

Base dati



ViaggioTreno: API ufficiali di Trenitalia

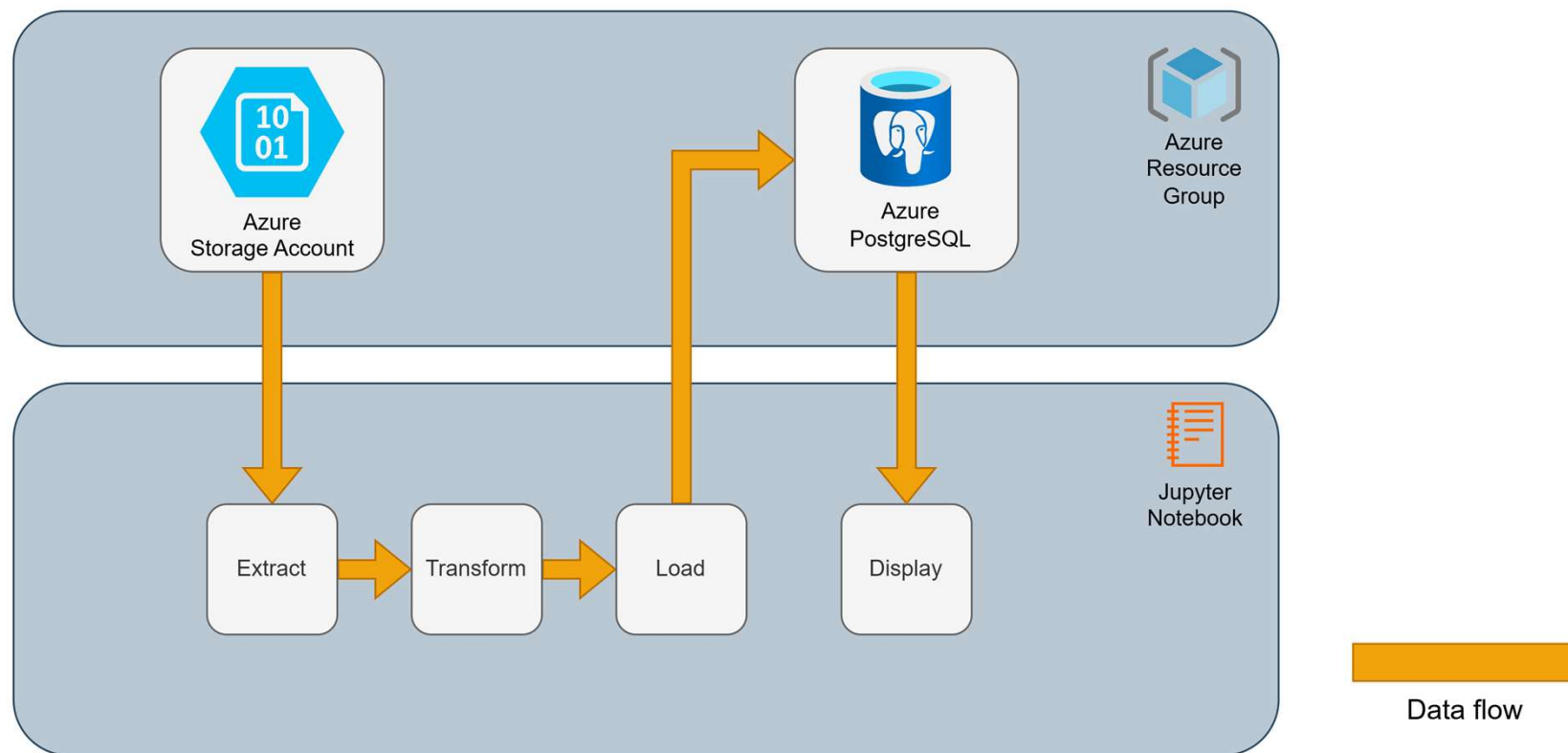
- PRO:
 - Dati ufficiali Trenitalia
 - Altissimo numero di informazioni (tempi, meteo, avvisi)
- CONTRO:
 - Intuitività delle API molto bassa
 - Non è possibile interrogare lo storico dati



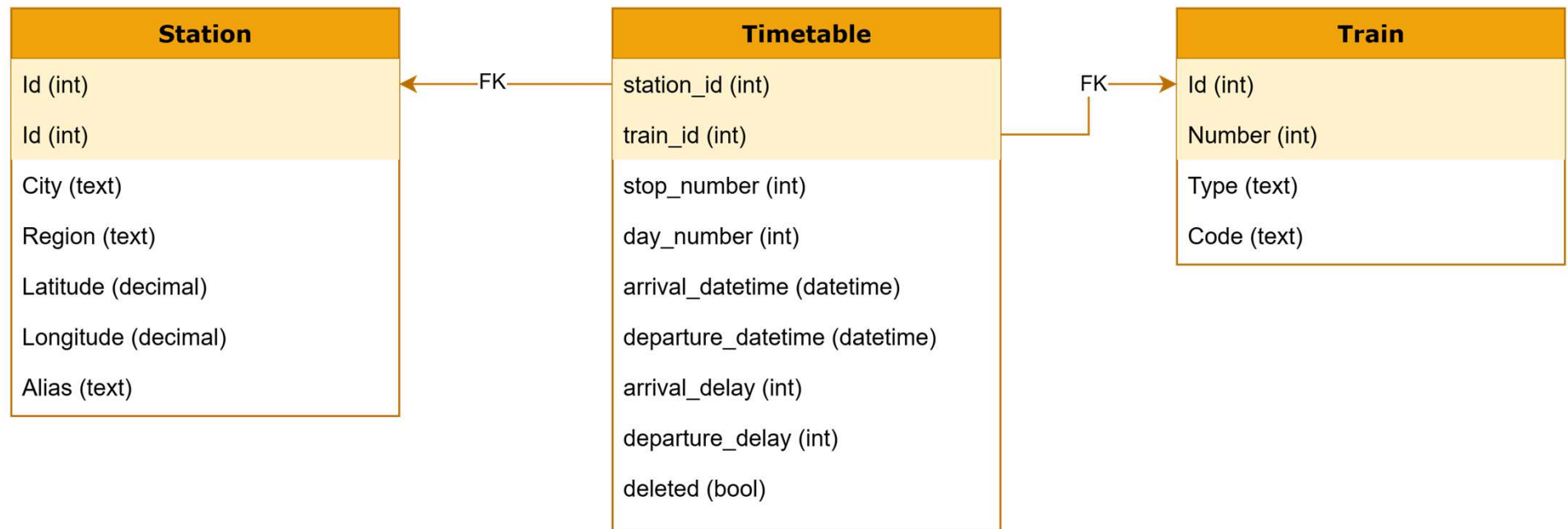
TrainStats: File JSON esposti pubblicamente su DropBox

- PRO:
 - Grande base dati (Luglio 2024 – Giugno 2025)
 - Poca frammentazione dei dati, un file JSON per giorno
- CONTRO:
 - Dati NON normalizzati (es. nome stazioni NON univoche «VENEZIA S.LUCIA», «VENEZIA SANTA LUCIA»)
 - Necessario estrarre informazioni da un JSON complesso

Architettura

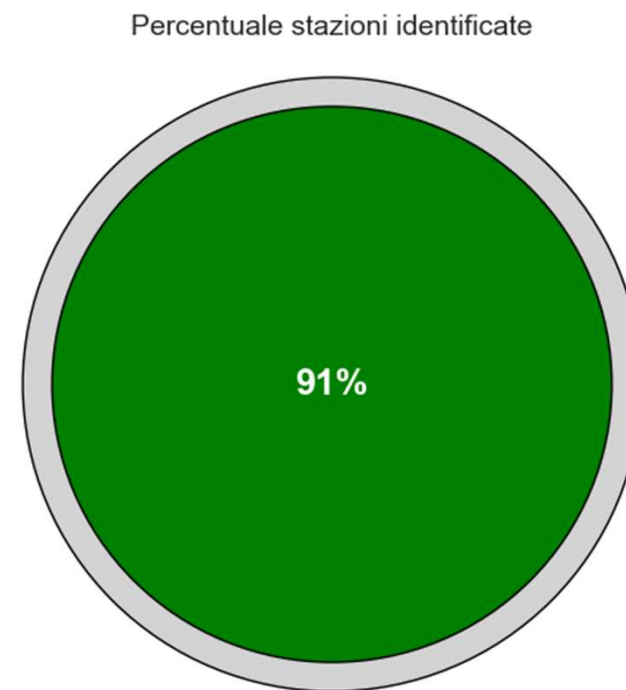
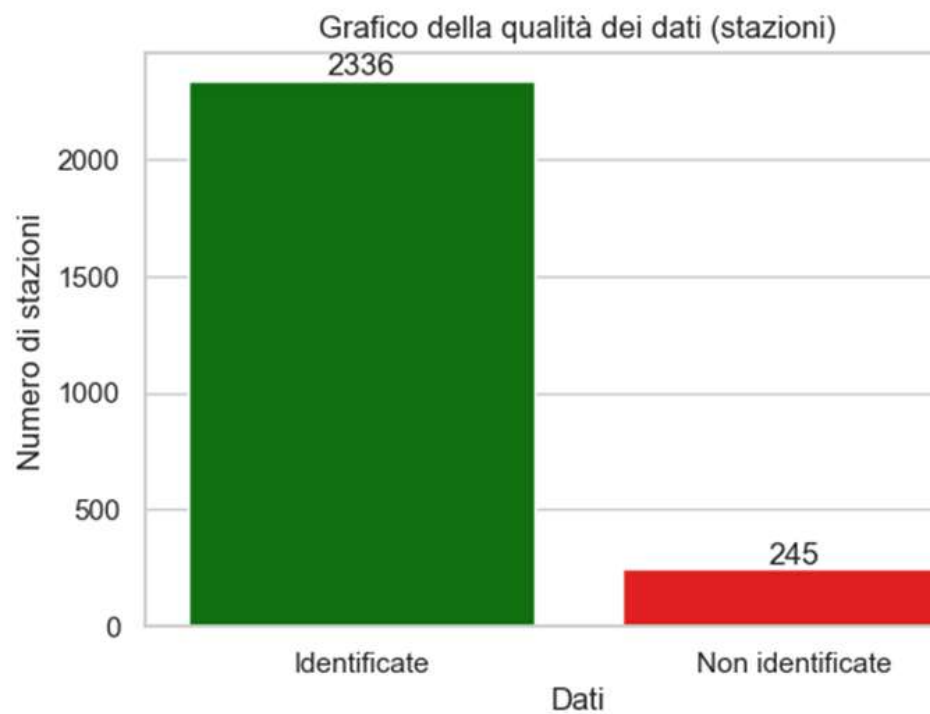


Struttura Database



Presentazione dati

Qualità dei dati: Stazioni identificate geograficamente

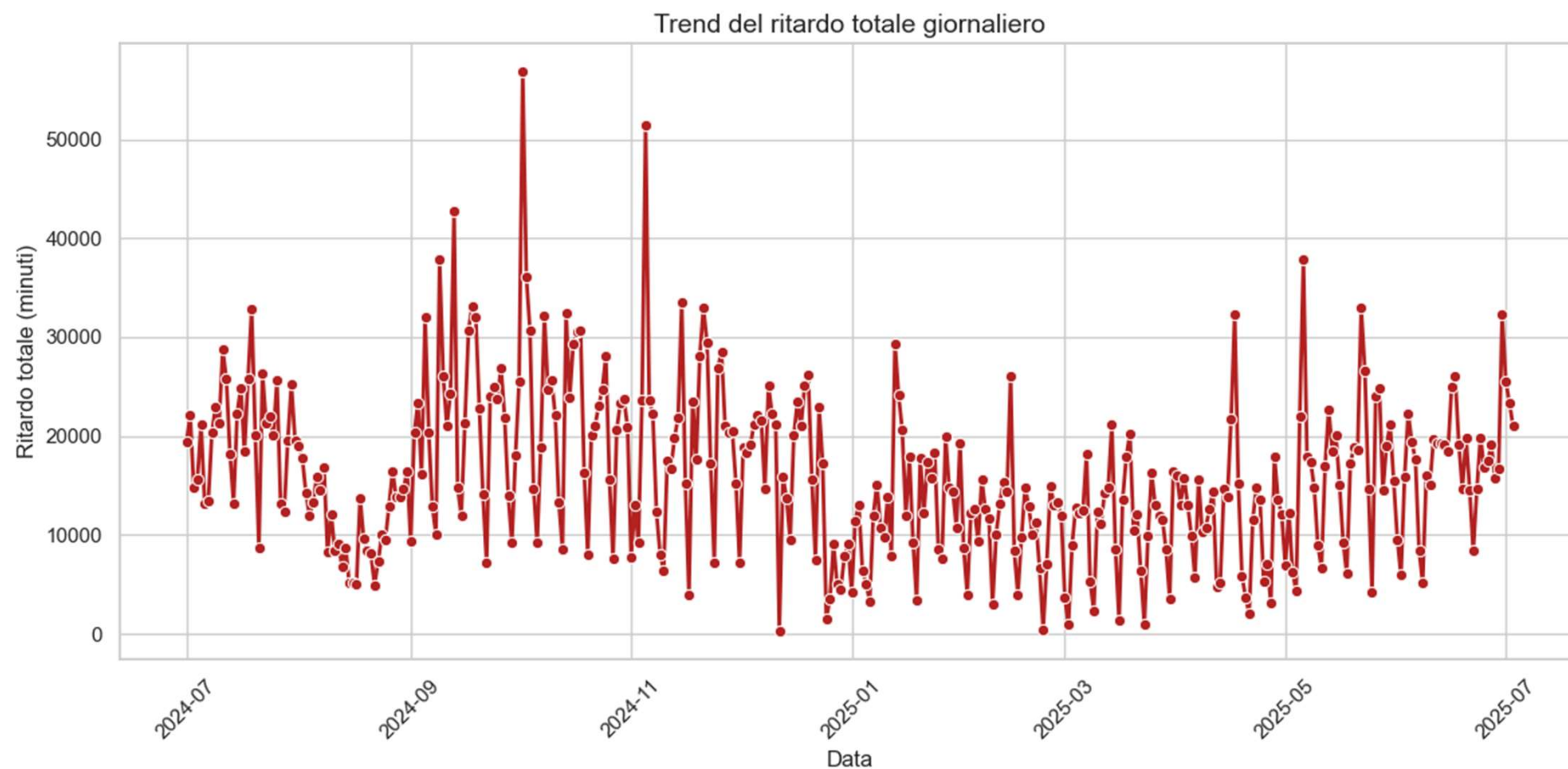


Distribuzione stazioni nel territorio italiano

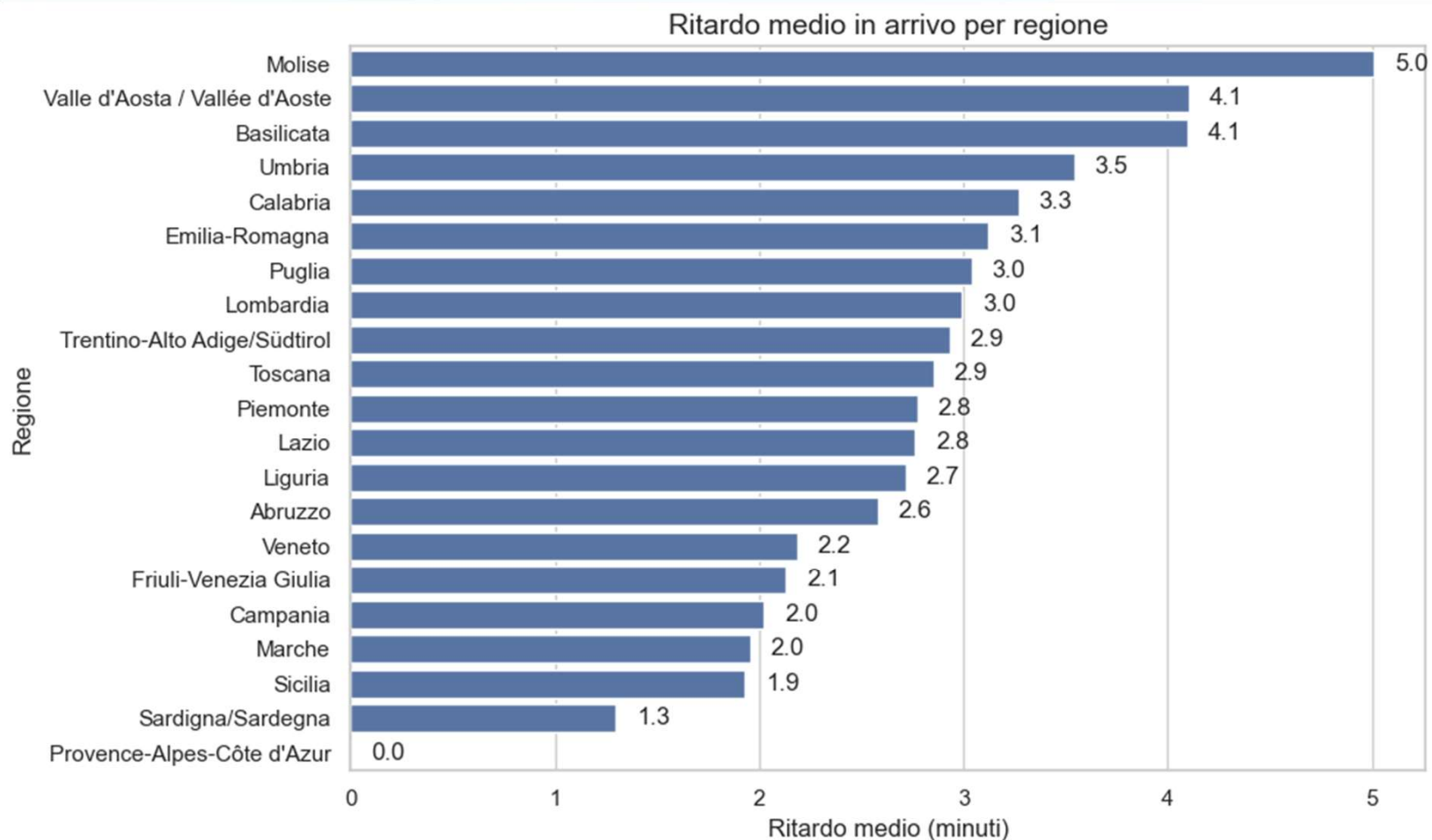
Stazioni ferroviarie in Italia



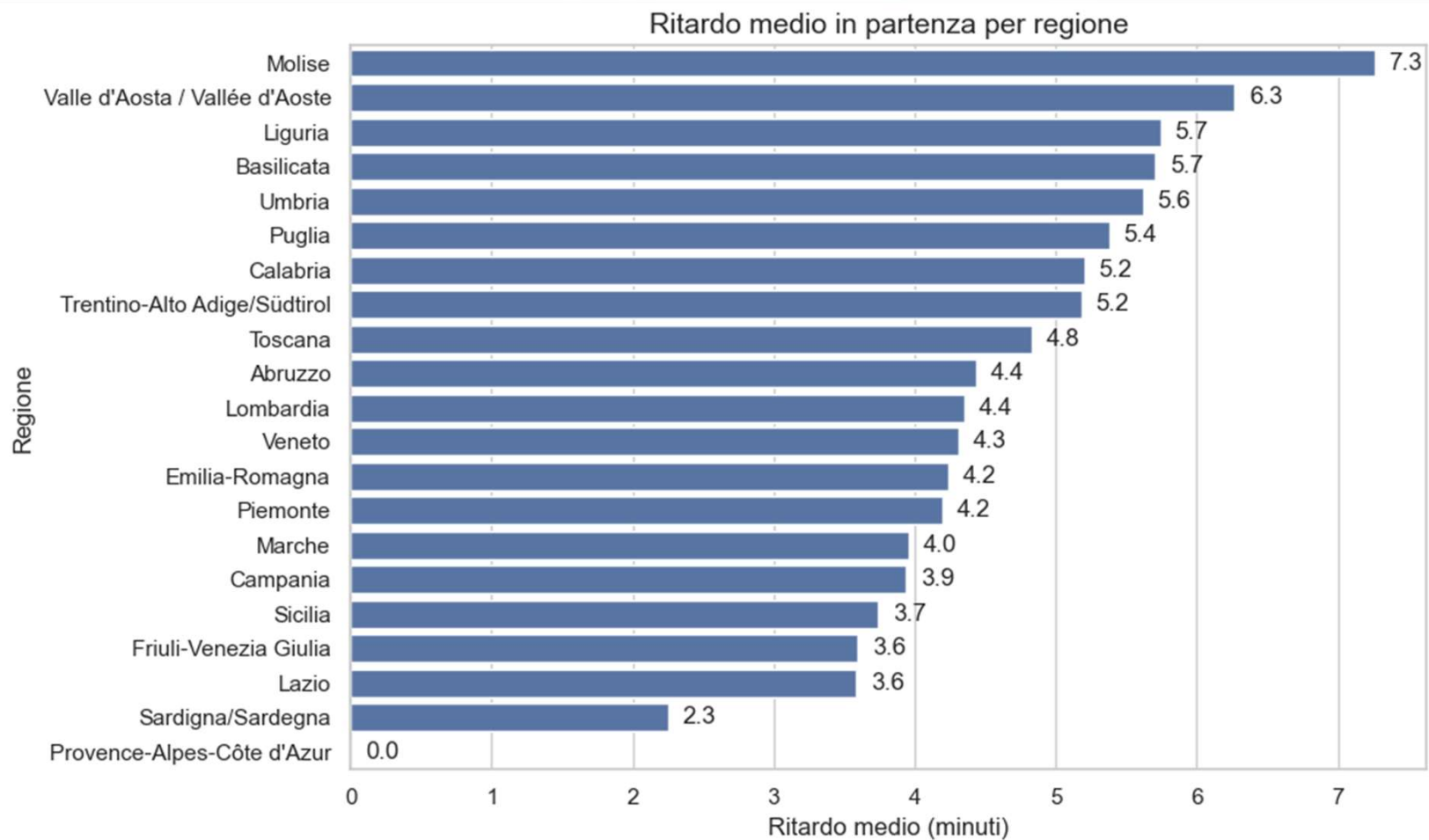
Trend ritardo nel periodo selezionato



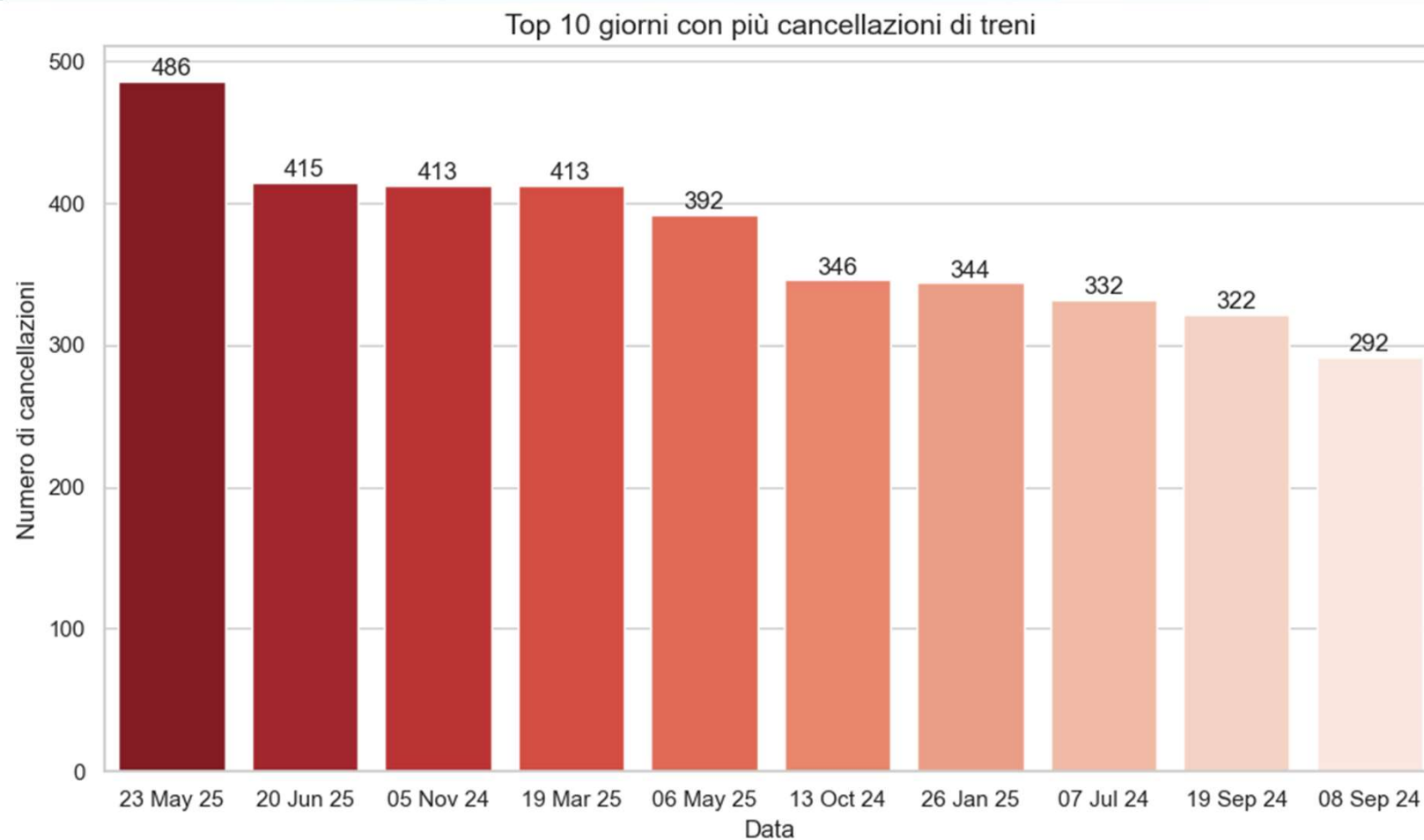
Ritardo medio arrive per regione



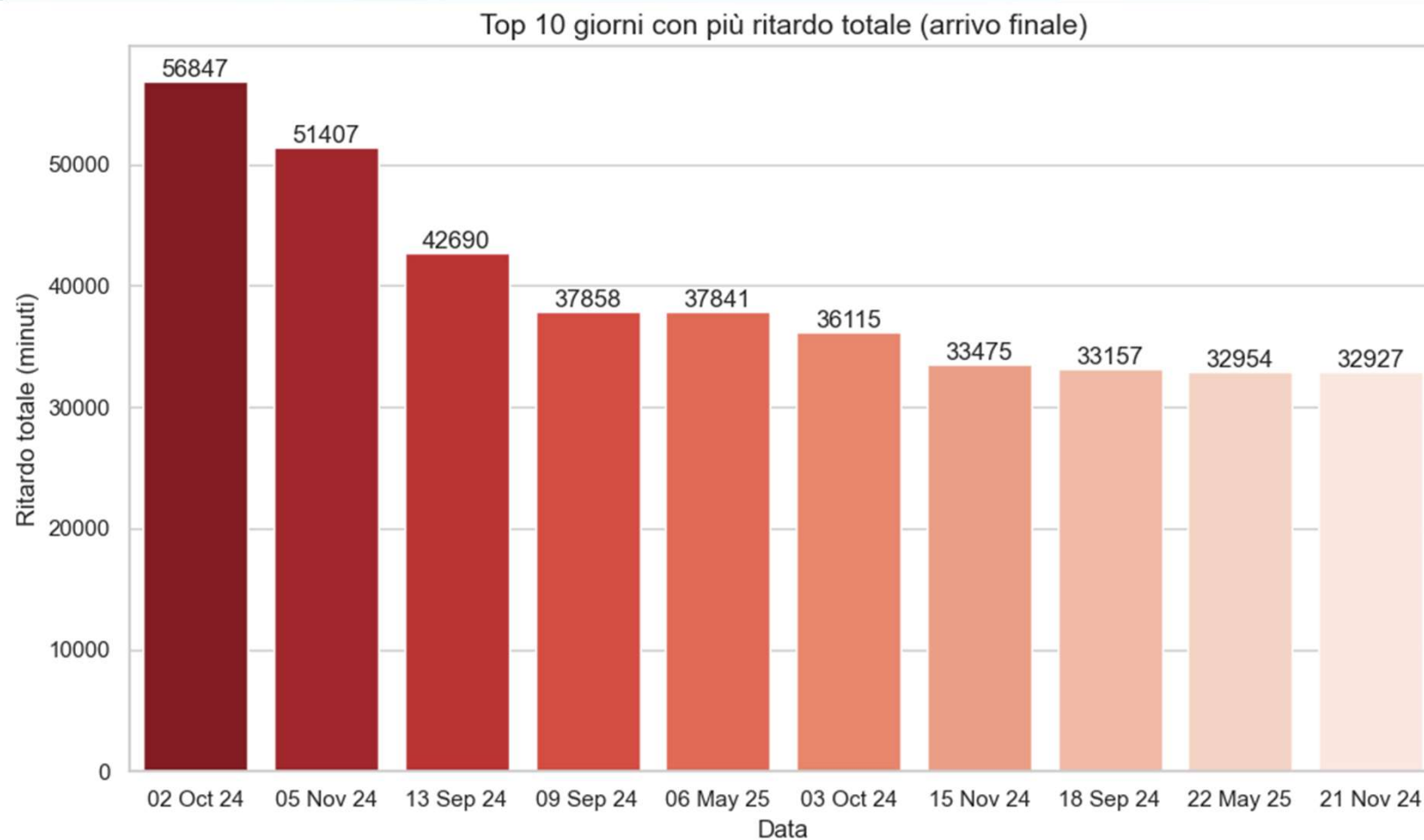
Ritardo medio partenze per regione



Giorni con maggior numero di cancellazioni



Giorni con maggior minuti di ritardo



Tratte con maggior minuti di ritardo medio (> 100 viaggi)

